PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 14/00, C12P 21/02, C12N 9/12, C07H 21/04

A1

(11) International Publication Number: WO 97/25345

(43) International Publication Date: 17 July 1997 (17.07.97)

(21) International Application Number: PCT/US97/00140

(22) International Filing Date: 3 January 1997 (03.01.97)

(30) Priority Data:
60/009,629 5 January 1996 (05.01.96) US

(71) Applicant *(for all designated States except US)*: ELI LILLY AND COMPANY [US/US]; Lilly Corporate Center, Indianapolis, IN 46285 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: KOVACEVIC, Steven [US/US]; 5620 Broadway, Indianapolis, IN 46220 (US). OTTO, Keith, A. [US/US]; 5245 Bobs Court, Greenwood, IN 46143 (US). RAO, Ramachandra, N. [US/US]; 9110 Misty Lake Circle, Indianapolis, IN 46260 (US).

(74) Agents: BLALOCK, Donna, K. et al.; Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description.
Date of receipt by the International Bureau:
27 March 1997 (27.03.97)

(54) Title: FUSION PROTEINS COMPRISING CELL CYCLE REGULATORY PROTEINS

(57) Abstract

The present invention provides novel fusion proteins comprising cyclins and CDKs. A preferred embodiment of the invention provides fusion proteins comprising human cyclin D1 and human CDK4. The fusion proteins of the invention optionally contain modifications, which facilitate their purification. Addition of histidine residues to selected constructs allows purification via immobilized metal affinity chromatography. Antigenic determinants allowing monoclonal antibody-based affinity chromatography purification are provided in selected embodiments of the invention. Protease cleavage sites are incorporated in selected constructs to allow cleavage of the regions incorporated in the cyclin-CDK fusion proteins for purification. Additional modifications which facilitate purification include strepavadin binding domains and antigenic determinants for antibody affinity chromatography.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

Title

FUSION PROTEINS COMPRISING CELL CYCLE REGULATORY PROTEINS

Background of the Invention

The pivotal roles which cyclins and cyclin dependent kinases play in cell cycle regulation is well established. The initial interest in cyclins resulted from observations that this family of molecules accumulated and then disappeared at precise points in the cell cycles of embryonic cells. Evans, T. *et al.*, Cell **33**, 389-396. (1983). Cyclin-dependent protein kinase (CDK) activation requires cyclin binding and phosphorylation of a threonine residue by the CDK-activating kinase, CAK. Several recent review articles (Norbury, c. and Nurse, P. A. Rev. Biochem. **61**, 441-470 (1992); Nasmyth, K. Curr. Opin. Cell Biol. **5**, 166-179 (1993) and Sherr, C. J. Cell **73**, 1059-1065 (1993)) detail the regulatory roles which the cyclins and the cyclin dependent kinases play in cell cycle progression.

The criticality of proper cell cycle regulation is intuitive. Disruption of cell cycle regulation leads to uncontrolled cell division. Appreciation of the important roles which cyclins and cyclin dependent kinases play in cell cycle regulation has focused intense research efforts aimed at better understanding cell cycle regulation and then exploiting this knowledge for discovery and development of oncoltyics.

Exploitation of the current knowledge regarding cyclins and CDKs requires experiments involving the addition of appropriate amounts of cyclins and CDKs to allow formation of the desired cyclin-CDK complex for phosphorylation of the conserved threonine residue of the CDK prior to attempting to modulate CDK-mediated phosphorylation of the retinoblastoma protein, Rb. The stochiometric problems inherent in such complicated experimental designs are substantial. The present invention addresses this problem by providing fusion proteins comprising cyclins and CDK4. The biological activities of

these fusion proteins eliminates the stochiometry related problems.

Summary of the Invention

The present invention provides novel fusion proteins comprising cyclins and CDKs. A preferred embodiement of the invention provides fusion proteins comprising human cyclin D1 and human CDK4. The fusion proteins of the invention optionally contain modifications, which facilitate their purification. Addition of histidine residues to selected constructs allows purification via immobilized metal affinity chromatography. Antigenic determinants allowing monoclonal antibody-based affinity chromatography purification are provided in selected embodiments of the invention. Proetease cleavage sites are incorporated in selected constructs to allow cleavage of the regions incorporated in the cyclin-CDK fusion proteins for purification. Additional modifications which facilitate purification include strepavadin binding domains and antigenic determinants for antibody affinity chromatography.

Brief Description of the Figures

Figure 1 is a restriction site and function map of plasmid pK415.

Figure 2 is a restriction site and function map of plasmid pK485.

Figure 3 is a restriction site and function map of plasmid pK480.

Detailed Description of the Invention

The fusion proteins of the present invention comprise cyclins and CDKs linked via various peptide spacers and optionally contain amino acid sequences, which are incorporated to facilitate purification.

The DNA sequence **(SEQ ID NO:1)** encoding a preferred embodiment of the present invention is provided below.

1 GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTTAA TGTCATGATA

51 ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG

101 AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA

151 TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT

201 ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT

251 TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG

301 CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC

351 AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT

401 GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTATTGACG

451 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG

501 GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT

551 AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA

601 ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG

651 CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT

701 GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGTAGCAA

751 TGGCAACAAC GTTGCGCAAA CTATTAACTG GCGAACTACT TACTCTAGCT

801 TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC

851 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG

901 GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT

951 GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC

1001 TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA

1051 AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT

1101 TTAAAACTTC ATTTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA

1151 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT

1201 CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTTCTG

1251 CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT

1301 TGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACTGGCT

1351 TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA

1401 GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT

1451 AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG

1501 GGTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA

1551 ACGGGGGGTT CGTGCACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA

1601 ACTGAGATAC CTACAGCGTG AGCATTGAGA AAGCGCCACG CTTCCCGAAG

1651 GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG

1701 CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT

1751 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTTGTGA TGCTCGTCAG

1801 GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGGCCTT TTTACGGTTC

1851 CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTTCCTG CGTTATCCCC

1901 TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC

1951 GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA

2001 GAGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC CGATTCATTA

2051 ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA

2101 ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC

2151 TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT

2201 TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCTT ACGGCGCGCC

2251 GCCGCCACCA TGGCGGAGGA GCAGAAGCTG ATATCCGAGG AGGACCTGCT

2301 GCTAGCAATG GAACACCAGC TCCTGTGCTG CGAAGTGGAA ACCATCCGCC

2351 GCGCGTACCC CGATGCCAAC CTCCTCAACG ACCGGGTGCT GCGGGCCATG

2401 CTGAAGGCGG AGGAGACCTG CGCGCCCTCG GTGTCCTACT TCAAATGTGT

2451 GCAAAAGGAG GTCCTGCCGT CCATGCGGAA GATCGTCGCC ACCTGGATGC

2501 TGGAGGTCTG CGAGGAACAG AAGTGCGAGG AGGAGGTCTT CCCGCTGGCC

2551 ATGAACTACC TGGACCGCTT CCTGTCGCTG GAGCCCGTGA AAAAGAGCCG

2601 CCTGCAGCTG CTGGGGGCCA CTTGCATGTT CGTGGCCTCT AAGATGAAGG

2651 AGACCATCCC CCTGACGGCC GAGAAGCTGT GCATCTACAC CGACAACTCC

2701 ATCCGGCCCG AGGAGCTGCT GCAAATGGAG CTGCTCCTGG TGAACAAGCT

2751 CAAGTGGAAC CTGGCCGCAA TGACCCCGCA CGATTTCATT GAACACTTCC

2801 TCTCCAAAAT GCCAGAGGCG GAGGAGAACA AACAGATCAT CCGCAAACAC

2851 GCGCAGACCT TCGTTGCCCT CTGTGCCACA GATGTGAAGT TCATTTCCAA

2901 TCCGCCCTCC ATGGTGGCAG CGGGGAGCGT GGTGGCCGCA GTGCAAGGCC

2951 TGAACCTGAG GAGCCCCAAC AACTTCCTGT CCTACTACCG CCTCACACGC

3001 TTCCTCTCCA GAGTGATCAA GTGTGACCCA GACTGCCTCC GGGCCTGCCA

3051 GGAGCAGATC GAAGCCCTGC TGGAGTCAAG CCTGCGCCAG GCCCAGCAGA

3101 ACATGGACCC CAAGGCCGCC GAGGAGGAGG AGGAGGAAGA GGAGGAAGAG

3151 GAGGTGGACC TGGCTTGCAC ACCCACCGAC GTGCGGGACG TGGACATCGC

3201 ATCGAAGGGT GGTGGAGGTT CTGGAGGTGG AGGATCCGGT GGTGGAGGTT

3251 CGATGGCTAC CTCTCGATAT GAGCCAGTGG CTGAAATTGG TGTCGGTGCC

3301 TATGGGACAG TGTACAAGGC CCGTGATCCC CACAGTGGCC ACTTTGTGGC

3351 CCTCAAGAGT GTGAGAGTCC CCAATGGAGG AGGAGGTGGA GGAGGCCTTC

3401 CCATCAGCAC AGTTCGTGAG GTGGCTTTAC TGAGGCGACT GGAGGCTTTT

3451 GAGCATCCCA ATGTTGTCCG GCTGATGGAC GTCTGTGCCA CATCCCGAAC

3501 TGACCGGGAG ATCAAGGTAA CCCTGGTGTT TGAGCATGTA GACCAGGACC

3551 TAAGGACATA TCTGGACAAG GCACCCCCAC CAGGCTTGCC AGCCGAAACG

3601 ATCAAGGATC TGATGCGCCA GTTTCTAAGA GGCCTAGATT TCCTTCATGC

3651 CAATTGCATC GTTCACCGAG ATCTGAAGCC AGAGAACATT CTGGTGACAA

3701 GTGGTGGAAC AGTCAAGCTG GCTGACTTTG GCCTGGCCAG AATCTACAGC

3751 TACCAGATGG CACTTACACC CGTGGTTGTT ACACTCTGGT ACCGAGCTCC

3801 CGAAGTTCTT CTGCAGTCCA CATATGCAAC ACCTGTGGAC ATGTGGAGTG

3851 TTGGCTGTAT CTTTGCAGAG ATGTTTCGTC GAAAGCCTCT CTTCTGTGGA

3901 AACTCTGAAG CCGACCAGTT GGGCAAAATC TTTGACCTGA TTGGGCTGCC

3951 TCCAGAGGAT GACTGGCCTC GAGATGTATC CCTGCCCCGT GGAGCCTTTC

4001 CCCCCAGAGG GCCCCGCCCA GTGCAGTCGG TGGTACCTGA GATGGAGGAG

4051 TCGGGAGCAC AGCTGCTGCT GGAAATGCTG ACTTTTAACC CACACAAGCG

4101 AATCTCTGCC TTTCGAGCTC TGCAGCACTC TTATCTACAT AAGGATGAAG

4151 GTAATCCGGA GGGCGGCAGC GCTTGGCGCC ACCCACAGTT CGGTGGTTGA

4201 ATAAATAGAT GAATGACCTG CAGGTTCACT GGCCGTCGTT TTACAACGTC

4251 GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT

4301 CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC

4351 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCCTG ATGCGGTATT

4401 TTCTCCTTAC GCATCTGTGC GGTATTTCAC ACCGCATATG GTGCACTCTC

4451 AGTACAATCT GCTCTGATGC CGCATAGTTA AGCCAGCCCC GACACCCGCC

4501 AACACCCGCT GACGCGCCCT GACGGGCTTG TCTGCTCCCG GCATCCGCTT

4551 ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA GAGGTTTTCA

4601 CCGTCATCAC CGAAACGCGC GA

The polypeptide encoded by SEQ ID NO:1 is presented below as **SEQ ID NO:2**.

1 MTMITPSLRR AAATMAEEQK LISEEDLLLA MEHQLLCCEV ETIRRAYPDA

51 NLLNDRVLRA MLKAEETCAP SVSYFKCVQK EVLPSMRKIV ATWMLEVCEE

101 QKCEEEVFPL AMNYLDRFLS LEPVKKSRLQ LLGATCMFVA SKMKETIPLT

151 AEKLCIYTDN SIRPEELLQM ELLLVNKLKW NLAAMTPHDF IEHFLSKMPE

201 AEENKQIIRK HAQTFVALCA TDVKFISNPP SMVAAGSVVA AVQGLNLRSP

251 NNFLSYYRLT RFLSRVIKCD PDCLRACQEQ IEALLESSLR QAQQNMDPKA

301 AEEEEEEEEE EEVDLACTPT DVRDVDIASK GGGGSGGGGS GGGGSMATSR

351 YEPVAEIGVG AYGTVYKARD PHSGHFVALK SVRVPNGGGG GGGLPISTVR

401 EVALLRRLEA FEHPNVVRLM DVCATSRTDR EIKVTLVFEH VDQDLRTYLD

451 KAPPPGLPAE TIKDLMRQFL RGLDFLHANC IVHRDLKPEN ILVTSGGTVK

501 LADFGLARIY SYQMALTPVV VTLWYRAPEV LLQSTYATPV DMWSVGCIFA

551 EMFRRKPLFC GNSEADQLGK IFDLIGLPPE DDWPRDVSLP RGAFPPRGPR

601 PVQSVVPEME ESGAQLLLEM LTFNPHKRIS AFRALQHSYL HKDEGNPEGG

651 SAWRHPQFGG

The DNA sequence of **SEQ ID NO:1** is the preferred coding sequence for the polypeptide of **SEQ ID NO:2**. Numerous other DNA sequences will also encode the polypeptide of **SEQ ID NO:2** due to the degeneracy of the genetic code. All DNA sequences encoding the polypeptide of **SEQ ID NO:2** are

contemplated by the present invention and thus are within the scope of the present invention.

The DNA sequence of **SEQ ID NO:1** is a component of the plasmid K415. A restriction site and function map of plasmid K415 is provided in Figure 1. *E. coli* host cells transformed with K415 were deposited in the NRRL, Northern Regional Research Laboratory, 1815 North University Street, Peoria, Illinois 61604 on or before August 9, 1995 and will be available pursuant to Budapest Treaty requirements upon issuance of a patent in a Budapest signatory country. The NRRL accession number for E. coli/K415 is B-21490. The routine nature of culturing such organisms, preparing plasmids from the transformants, digesting the plasmids with appropriate restriction endonucleases and isolating the appropriate DNA fragment obviate the need or desirability of discussing these routine steps.

The distinct functional subcomponents of the polypeptide of **SEQ ID NO:2** are described by reference to the amino acid residue numbers provided in **SEQ ID NO:2**. Residues 18 through 27 comprise the epitope recognized by the monoclonal antibody designated myc. Residues 31 though 327 correspond to human cyclin D1. Residues 331 through 345 are an illustrative "linker" or polypedtide connector. The terms "linker", "polypeptide connector" and "hinge" are used interchangeably in describing the present invention and all three terms refer to the sequences of amino acids which are used to connect the cyclin and CDK components of the fusion proteins of the present invention. Residues 346 through 648 correspond to human CDK4. Residues 651 through 660 correspond to strepavadin and were engineered into the molecule to allow facile purification.

The polypeptide of **SEQ ID NO:2** has numerous components which allow great flexibility in purification, but are not required for the ultimate benefit provided by the present invention-a biologically active fusion protein comprising cyclin and CDK components. A most preferred aspect of this embodiment of the present invention is the

cyclin D1-linker-CDK4 component of the molecule. This most preferred aspect is provided below as **SEQ ID NO:3**.

41 MEHQLLCCEV ETIRRAYPDA

51 NLLNDRVLRA MLKAEETCAP SVSYFKCVQK EVLPSMRKIV ATWMLEVCEE

101 QKCEEEVFPL AMNYLDRFLS LEPVKKSRLQ LLGATCMFVA SKMKETIPLT

151 AEKLCIYTDN SIRPEELLQM ELLLVNKLKW NLAAMTPHDF IEHFLSKMPE

201 AEENKQIIRK HAQTFVALCA TDVKFISNPP SMVAAGSVVA AVQGLNLRSP

251 NNFLSYYRLT RFLSRVIKCD PDCLRACQEQ IEALLESSLR QAQQNMDPKA

301 AEEEEEEEEE EEVDLACTPT DVRDVDIASK GGGGSGGGGS GGGGSMATSR

351 YEPVAEIGVG AYGTVYKARD PHSGHFVALK SVRVPNGGGG GGGLPISTVR

401 EVALLRRLEA FEHPNVVRLM DVCATSRTDR EIKVTLVFEH VDQDLRTYLD

451 KAPPPGLPAE TIKDLMRQFL RGLDFLHANC IVHRDLKPEN ILVTSGGTVK

501 LADFGLARIY SYQMALTPVV VTLWYRAPEV LLQSTYATPV DMWSVGCIFA

551 EMFRRKPLFC GNSEADQLGK IFDLIGLPPE DDWPRDVSLP RGAFPPRGPR

601 PVQSVVPEME ESGAQLLLEM LTFNPHKRIS AFRALQHSYL HKDEGNPE

Biologically active fusion protein comprising a member of the cyclin family and the CDK family are further illustrated by the DNA sequence of **SEQ ID NO:4** and the corresponding polypeptide sequence, **SEQ ID NO:5**. **SEQ ID NO:4** is provided immediately below.

1 GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTTAA TGTCATGATA

51 ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG

101 AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA

151 TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT

201 ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT

251 TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG

301 CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC

351 AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT

401 GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTATTGACG

451 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG

501 GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT

551 AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA

601 ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG

651 CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT

701 GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGTAGCAA

751 TGGCAACAAC GTTGCGCAAA CTATTAACTG GCGAACTACT TACTCTAGCT

801 TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC

851 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG

901 GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT

951 GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC

1001 TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA

1051 AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT

1101 TTAAAACTTC ATTTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA

1151 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT

1201 CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTTCTG

1251 CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT

1301 TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACTGGCT

1351 TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA

1401 GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT

1451 AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG

1501 GGTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA

1551 ACGGGGGGTT CGTGCACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA

1601 ACTGAGATAC CTACAGCGTG AGCATTGAGA AAGCGCCACG CTTCCCGAAG

1651 GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG

1701 CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT

1751 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTTGTGA TGCTCGTCAG

1801 GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGGCCTT TTTACGGTTC

1851 CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTTCCTG CGTTATCCCC

1901 TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC

1951 GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA

2001 GAGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC CGATTCATTA

2051 ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA

2101 ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC

2151 TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT

2201 TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCTT ACGGCGCGCC

2251 GCCGCCACCA TGGCGCATCA TCATCATCAT CATGGAGGTG GAGGTTCGGA

2301 GCAGAAGCTT ATTTCCGAGG AGGATCTGCT GGTGCCACGC GGTTCCCTGC

2351 TAGCAATGGA ACACCAGCTC CTGTGCTGCG AAGTGGAAAC CATCCGCCGC

2401 GCGTACCCCG ATGCCAACCT CCTCAACGAC CGGGTGCTGC GGGCCATGCT

2451 AAAGGCGGAG GAGACCTGCG CGCCCTCGGT GTCCTACTTC AAATGTGTGC

2501 AAAAGGAGGT CCTGCCGTCC ATGCGGAAGA TCGTCGCCAC CTGGATGCTG

2551 GAGGTCTGCG AGGAACAGAA GTGCGAGGAG GAGGTCTTCC CGCTGGCCAT

2601 GAACTACCTG GACCGCTTCC TGTCGCTGGA GCCCGTGAAA AAGAGCCGCC

2651 TGCAGCTGCT GGGGGCCACT TGCATGTTCG TGGCCTCTAA GATGAAGGAG

2701 ACCATCCCCC TGACGGCCGA GAAGCTGTGC ATCTACACCG ACAACTCCAT

2751 CCGGCCCGAG GAGCTGCTGC AAATGGAGCT GCTCCTGGTG AACAAGCTCA

2801 AGTGGAACCT GGCCGCAATG ACCCCGCACG ATTTCATTGA ACACTTCCTC

2851 TCCAAAATGC CAGAGGCGGA GGAGAACAAA CAGATCATCC GCAAACACGC

2901 GCAGACCTTC GTTGCCCTCT GTGCCACAGA TGTGAAGTTC ATTTCCAATC

2951 CGCCCTCCAT GGTGGCAGCG GGGAGCGTGG TGGCCGCAGT GCAAGGCCTG

3001 AACCTGAGGA GCCCCAACAA CTTCCTGTCC TACTACCGCC TCACACGCTT

3051 CCTCTCCAGA GTGATCAAGT GTGACCCAGA CTGCCTCCGG GCCTGCCAGG

3101 AGCAGATCGA AGCCCTGCTG GAGTCAAGCC TGCGCCAGGC CCAGCAGAAC

3151 ATGGACCCCA AGGCCGCCGA GGAGGAGGAG GAGGAAGAGG AGGAAGAGGA

3201 GGTGGACCTG GCTTGCACAC CCACCGACGT GCGGGACGTG GACATCGCAT

3251 CGAAGGGTGG TGGAGGTTCT GGAGGTGGAG GATCCGGTGG TGGAGGTTCG

3301 ATGGCTACCT CTCGATATGA GCCAGTGGCT GAAATTGGTG TCGGTGCCTA

3351 TGGGACAGTG TACAAGGCCC GTGATCCCCA CAGTGGCCAC TTTGTGGCCC

3401 TCAAGAGTGT GAGAGTCCCC AATGGAGGAG GAGGTGGAGG AGGCCTTCCC

3451 ATCAGCACAG TTCGTGAGGT GGCTTTACTG AGGCGACTGG AGGCTTTTGA

3501 GCATCCCAAT GTTGTCCGGC TGATGGACGT CTGTGCCACA TCCCGAACTG

3551 ACCGGGAGAT CAAGGTAACC CTGGTGTTTG AGCATGTAGA CCAGGACCTA

3601 AGGACATATC TGGACAAGGC ACCCCCACCA GGCTTGCCAG CCGAAACGAT

3651 CAAGGATCTG ATGCGCCAGT TTCTAAGAGG CCTAGATTTC CTTCATGCCA

3701 ATTGCATCGT TCACCGAGAT CTGAAGCCAG AGAACATTCT GGTGACAAGT

3751 GGTGGAACAG TCAAGCTGGC TGACTTTGGC CTGGCCAGAA TCTACAGCTA

3801 CCAGATGGCA CTTACACCCG TGGTTGTTAC ACTCTGGTAC CGAGCTCCCG

3851 AAGTTCTTCT GCAGTCCACA TATGCAACAC CTGTGGACAT GTGGAGTGTT

3901 GGCTGTATCT TTGCAGAGAT GTTTCGTCGA AAGCCTCTCT TCTGTGGAAA

3951 CTCTGAAGCC GACCAGTTGG GCAAAATCTT TGACCTGATT GGGCTGCCTC

4001 CAGAGGATGA CTGGCCTCGA GATGTATCCC TGCCCCGTGG AGCCTTTCCC

4051 CCCAGAGGGC CCCGCCCAGT GCAGTCGGTG GTACCTGAGA TGGAGGAGTC

4101 GGGAGCACAG CTGCTGCTGG AAATGCTGAC TTTTAACCCA CACAAGCGAA

4151 TCTCTGCCTT TCGAGCTCTG CAGCACTCTT ATCTACATAA GGATGAAGGT

4201 AATCCGGAGG GCGGCAGCGC TTGGCGCCAC CCACAGTTCG GTGGTTGAAT

4251 AAATAGATGA ATGACCTGCA GGTGCACTCT CAGTACAATC TGCTCTGATG

4301 CCGCATAGTT AAGCCAGCCC CGACACCCGC CAACACCCGC TGACGCGCCC

4351 TGACGGGCTT GTCTGCTCCC GGCATCCGCT TACAGACAAG CTGTGACCGT

4401 CTCCGGGAGC TGCATGTGTC AGAGGTTTTC ACCGTCATCA CCGAAACGCG

4451 CGA

The polypeptide encoded by the sequence of **SEQ ID NO:4** is provided below as **SEQ ID NO:5**.

1 MAHHHHHHGG GGSEQKLISE EDLLVPRGSL LAMEHQLLCC EVETIRRAYP

51 DANLLNDRVL RAMLKAEETC APSVSYFKCV QKEVLPSMRK IVATWMLEVC

101 EEQKCEEEVF PLAMNYLDRF LSLEPVKKSR LQLLGATCMF VASKMKETIP

151 LTAEKLCIYT DNSIRPEELL QMELLLVNKL KWNLAAMTPH DFIEHFLSKM

201 PEAEENKQII RKHAQTFVAL CATDVKFISN PPSMVAAGSV VAAVQGLNLR

251 SPNNFLSYYR LTRFLSRVIK CDPDCLRACQ EQIEALLESS LRQAQQNMDP

301 KAAEEEEEEE EEEEVDLACT PTDVRDVDIA SKGGGGSGGG GSGGGGSMAT

351 SRYEPVAEIG VGAYGTVYKA RDPHSGHFVA LKSVRVPNGG GGGGGLPIST

401 VREVALLRRL EAFEHPNVVR LMDVCATSRT DREIKVTLVF EHVDQDLRTY

451 LDKAPPPGLP AETIKDLMRQ FLRGLDFLHA NCIVHRDLKP ENILVTSGGT

501 VKLADFGLAR IYSYQMALTP VVVTLWYRAP EVLLQSTYAT PVDMWSVGCI

551 FAEMFRRKPL FCGNSEADQL GKIFDLIGLP PEDDWPRDVS LPRGAFPPRG

601 PRPVQSVVPE MEESGAQLLL EMLTFNPHKR ISAFRALQHS YLHKDEGNPE

651 GGSAWRHPQF GG

The DNA sequence of **SEQ ID NO:4** is the preferred coding sequence for the polypeptide of **SEQ ID NO:5**. Numerous other DNA sequences will also encode the polypeptide of **SEQ ID NO:4** due to the degeneracy of the genetic code. All DNA sequences encoding the polypeptide of **SEQ ID NO:5** are contemplated by the present invention and thus are within the scope of the present invention.

The DNA sequence of **SEQ ID NO:4** is a component of the plasmid K485. A restriction site and function map of plasmid K485 is provided in **Figure 2**. *E. coli* host cells transformed with K485 were deposited in the NRRL, Northern Regional Research Laboratory, 1815 North University Street, Peoria, Illinois 61604 on or before August 9, 1995 and will

be available pursuant to Budapest Treaty requirements upon issuance of a patent in a Budapest signatory country. The NRRL accession number for *E. coli*/K485 is B-21492. The routine nature of culturing such organisms, preparing plasmids from the transformants, digesting the plasmids with appropriate restriction endonucleases and isolating the appropriate DNA fragment obviate the need or desirability of discussing these routine steps.

The DNA sequence of Sequence ID 4 and the polypeptide encoded thereby comprise human cyclin D1 and human CDK4 which are joined by a polypeptide linker. The distinct functional subcomponents of the polypeptide of **SEQ ID NO:5** are described by reference to the amino acid residue numbers provided in **SEQ ID NO:5**. Amino acid residues 2 through 8 are Histidine residues which were incorporated to allow immobilized metal affinity chromatography purification. Residues 14 through 23 contain the antigenic determinant recognized by the myc monoclonal antibody and thereby allow myc monoclonal antibody based affinity purification. Residues 24 through 28 contain a thrombin cleavage site and were engineered into the polypeptide of SEQ ID NO:5 to allow cleavage of the molecule on the amino side of the human cyclin D1 component. Residues 43 through 329 correspond to human cyclin D1. Residues 333 through 347 are the polypeptide linker used to join the human cyclin D1 and human CDK4 components of the molecule. Residues 348 through 650 correspond to human CDK4. Residues 653 through 662 were engineered into the molecule to provide a sequence which binds to paramagnetic streptavadin beads and thus allows facile purification of the molecule.

The present invention also provides the DNA sequence of **SEQ ID NO:6**, which is presented below.

1 GACGAAAGGG CCTCGTGATA CGCCTATTTT TATAGGTTAA TGTCATGATA

51 ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG

```
101  AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA

151  TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT

201  ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT

251  TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG

301  CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC

351  AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT

401  GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC CGTATTGACG

451  CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG

501  GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT

551  AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA

601  ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG

651  CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT

701  GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGTAGCAA

751  TGGCAACAAC GTTGCGCAAA CTATTAACTG GCGAACTACT TACTCTAGCT

801  TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC

851  ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG

901  GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT

951  GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC

1001 TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA
```

1051 AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT

1101 TTAAAACTTC ATTTTTAATT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA

1151 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT

1201 CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTTCTG

1251 CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT

1301 TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACTGGCT

1351 TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA

1401 GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT

1451 AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG

1501 GGTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA

1551 ACGGGGGGTT CGTGCACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA

1601 ACTGAGATAC CTACAGCGTG AGCATTGAGA AAGCGCCACG CTTCCCGAAG

1651 GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG

1701 CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT

1751 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTTGTGA TGCTCGTCAG

1801 GGGGGCGGAG CCTATGGAAA AACGCCAGCA ACGCGGCCTT TTTACGGTTC

1851 CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTTCCTG CGTTATCCCC

1901 TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC

1951 GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA

2001 GAGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC CGATTCATTA

2051 ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA

2101 ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC

2151 TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT

2201 TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCTT ACGGCGCGCC

2251 GCCGCCACCA TGGCGCATCA TCATCATCAT CATGGAGGTG GAGGTTCGGA

2301 GCAGAAGCTT ATTTCCGAGG AGGATCTGCT GGTGCCACGC GGTTCCCTGC

2351 TAGCAATGGA ACACCAGCTC CTGTGCTGCG AAGTGGAAAC CATCCGCCGC

2401 GCGTACCCCG ATGCCAACCT CCTCAACGAC CGGGTGCTGC GGGCCATGCT

2451 AAAGGCGGAG GAGACCTGCG CGCCCTCGGT GTCCTACTTC AAATGTGTGC

2501 AAAAGGAGGT CCTGCCGTCC ATGCGGAAGA TCGTCGCCAC CTGGATGCTG

2551 GAGGTCTGCG AGGAACAGAA GTGCGAGGAG GAGGTCTTCC CGCTGGCCAT

2601 GAACTACCTG GACCGCTTCC TGTCGCTGGA GCCCGTGAAA AAGAGCCGCC

2651 TGCAGCTGCT GGGGGCCACT TGCATGTTCG TGGCCTCTAA GATGAAGGAG

2701 ACCATCCCCC TGACGGCCGA GAAGCTGTGC ATCTACACCG ACAACTCCAT

2751 CCGGCCCGAG GAGCTGCTGC AAATGGAGCT GCTCCTGGTG AACAAGCTCA

2801 AGTGGAACCT GGCCGCAATG ACCCCGCACG ATTTCATTGA ACACTTCCTC

2851 TCCAAAATGC CAGAGGCGGA GGAGAACAAA CAGATCATCC GCAAACACGC

2901 GCAGACCTTC GTTGCCCTCT GTGCCACAGA TGTGAAGTTC ATTTCCAATC

2951 CGCCCTCCAT GGTGGCAGCG GGGAGCGTGG TGGCCGCAGT GCAAGGCCTG

3001 AACCTGAGGA GCCCCAACAA CTTCCTGTCC TACTACCGCC TCACACGCTT

3051 CCTCTCCAGA GTGATCAAGT GTGACCCAGA CTGCCTCCGG GCCTGCCAGG

3101 AGCAGATCGA AGCCCTGCTG GAGTCAAGCC TGCGCCAGGC CCAGCAGAAC

3151 ATGGACCCCA AGGCCGCCGA GGAGGAGGAG GAGGAAGAGG AGGAAGAGGA

3201 GGTGGACCTG GCTTGCACAC CCACCGACGT GCGGGACGTG GACATCGCAT

3251 CGATGGGTGG AGGTTCTGGT GGAGGTTCTG GTGGAGGTTC TGGTGGAGGT

3301 TCTGGTGGAG GTTCTGGTGG AGGTTCTGGC TTAAGTTCGA AGGGTGGTGG

3351 AGGTTCTGGA GGTGGAGGAT CCGGTGGTGG AGGTTCGATG GCTACCTCTC

3401 GATATGAGCC AGTGGCTGAA ATTGGTGTCG GTGCCTATGG GACAGTGTAC

3451 AAGGCCCGTG ATCCCCACAG TGGCCACTTT GTGGCCCTCA AGAGTGTGAG

3501 AGTCCCCAAT GGAGGAGGAG GTGGAGGAGG CCTTCCCATC AGCACAGTTC

3551 GTGAGGTGGC TTTACTGAGG CGACTGGAGG CTTTTGAGCA TCCCAATGTT

3601 GTCCGGCTGA TGGACGTCTG TGCCACATCC CGAACTGACC GGGAGATCAA

3651 GGTAACCCTG GTGTTTGAGC ATGTAGACCA GGACCTAAGG ACATATCTGG

3701 ACAAGGCACC CCCACCAGGC TTGCCAGCCG AAACGATCAA GGATCTGATG

3751 CGCCAGTTTC TAAGAGGCCT AGATTTCCTT CATGCCAATT GCATCGTTCA

3801 CCGAGATCTG AAGCCAGAGA ACATTCTGGT GACAAGTGGT GGAACAGTCA

3851 AGCTGGCTGA CTTTGGCCTG GCCAGAATCT ACAGCTACCA GATGGCACTT

3901 ACACCCGTGG TTGTTACACT CTGGTACCGA GCTCCCGAAG TTCTTCTGCA

3951 GTCCACATAT GCAACACCTG TGGACATGTG GAGTGTTGGC TGTATCTTTG

4001 CAGAGATGTT TCGTCGAAAG CCTCTCTTCT GTGGAAACTC TGAAGCCGAC

4051 CAGTTGGGCA AAATCTTTGA CCTGATTGGG CTGCCTCCAG AGGATGACTG

4101 GCCTCGAGAT GTATCCCTGC CCCGTGGAGC CTTTCCCCCC AGAGGGCCCC

4151 GCCCAGTGCA GTCGGTGGTA CCTGAGATGG AGGAGTCGGG AGCACAGCTG

4201 CTGCTGGAAA TGCTGACTTT TAACCCACAC AAGCGAATCT CTGCCTTTCG

4251 AGCTCTGCAG CACTCTTATC TACATAAGGA TGAAGGTAAT CCGGAGGGCG

4301 GCAGCGCTTG GCGCCACCCA CAGTTCGGTG GTTGAATAAA TAGATGAATG

4351 ACCTGCAGGT GCACTCTCAG TACAATCTGC TCTGATGCCG CATAGTTAAG

4401 CCAGCCCCGA CACCCGCCAA CACCCGCTGA CGCGCCCTGA CGGGCTTGTC

4451 TGCTCCCGGC ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC

4501 ATGTGTCAGA GGTTTTCACC GTCATCACCG AAACGCGCGA

The polypeptide encoded by **SEQ ID NO:6** is presented below as **SEQ ID NO:7**.

1 MTMITPSLRR AAATMAHHHH HHGGGGSEQK LISEEDLLVP RGSLLAMEHQ

51 LLCCEVETIR RAYPDANLLN DRVLRAMLKA EETCAPSVSY FKCVQKEVLP

101 SMRKIVATWM LEVCEEQKCE EEVFPLAMNY LDRFLSLEPV KKSRLQLLGA

151 TCMFVASKMK ETIPLTAEKL CIYTDNSIRP EELLQMELLL VNKLKWNLAA

201 MTPHDFIEHF LSKMPEAEEN KQIIRKHAQT FVALCATDVK FISNPPSMVA

251 AGSVVAAVQG LNLRSPNNFL SYYRLTRFLS RVIKCDPDCL RACQEQIEAL

301 LESSLRQAQQ NMDPKAAEEE EEEEEEEVD LACTPTDVRD VDIASMGGGS

351 GGGSGGGSGG GSGGGSGGGS GLSSKGGGGS GGGGSGGGGS MATSRYEPVA

401 EIGVGAYGTV YKARDPHSGH FVALKSVRVP NGGGGGGGLP ISTVREVALL

451 RRLEAFEHPN VVRLMDVCAT SRTDREIKVT LVFEHVDQDL RTYLDKAPPP

501 GLPAETIKDL MRQFLRGLDF LHANCIVHRD LKPENILVTS GGTVKLADFG

551 LARIYSYQMA LTPVVVTLWY RAPEVLLQST YATPVDMWSV GCIFAEMFRR

601 KPLFCGNSEA DQLGKIFDLI GLPPEDDWPR DVSLPRGAFP PRGPRPVQSV

651 VPEMEESGAQ LLLEMLTFNP HKRISAFRAL QHSYLHKDEG NPEGGSAWRH

701 PQFGG

The DNA sequence of **SEQ ID NO:6** is the preferred coding sequence for the polypeptide of **SEQ ID NO:7**. Numerous other DNA sequences will also encode the polypeptide of **SEQ ID NO:6** due to the degeneracy of the genetic code. All DNA sequences encoding the polypeptide of **SEQ ID NO:7** are contemplated by the present invention and thus are within the scope of the present invention.

The DNA sequence of **SEQ ID NO:6** is a component of the plasmid K480. A restriction site and function map of plasmid K480 is provided in **Figure 3**. *E. coli* host cells transformed with K480 were deposited in the NRRL, Northern

Regional Research Laboratory, 1815 North University Street, Peoria, Illinois 61604 on or before August 9, 1995 and will be available pursuant to Budapest Treaty requirements upon issuance of a patent in a Budapest signatory country. The NRRL accession number for *E. coli*/K480 is B-21491. The routine nature of culturing such organisms, preparing plasmids from the transformants, digesting the plasmids with appropriate restriction endonucleases and isolating the appropriate DNA fragment obviate the need or desirability of discussing these routine steps.

The DNA sequence of **SEQ ID NO:6** and the polypeptide encoded thereby comprise human cyclin D1 and human CDK4 which are joined by a polypeptide linker. The distinct functional subcomponents of the polypeptide of **SEQ ID NO:7** are described by reference to the amino acid residue numbers provided in **SEQ ID NO:7**. Amino acid residues 17 through 22 are Histidine residues which were incorporated to allow immobilized metal affinity chromatography purification. Residues 28 through 37 contain the antigenic determinant recognized by the myc monoclonal antibody and thereby allow myc monoclonal antibody based affinity purification. Residues 38 through 43 contain a thrombin cleavage site and were engineered into the polypeptide of Sequence ID 7 to allow cleavage of the molecule on the amino side of the human cyclin D1 component. Residues 47 through 343 correspond to human cyclin D1. Residues 347 through 390 are the polypeptide linker used to join the human cyclin D1 and human CDK4 components of the molecule. Residues 391 through 693 correspond to human CDK4. Residues 696 through 705 were engineered into the molecule to provide a sequence which binds to paramagnetic streptavadin beads and thus allows facile purification of the molecule.

The molecule of **SEQ ID NO:7** shares several features with the molecules of **SEQ ID Nos:2 and 5**. The polypeptide linker which joins the human cyclin D1 and the human CDK4 portions of the molecule of **SEQ ID NO:7** is substantially different from the polypeptide linkers of the molecules of **SEQ ID Nos: 2 and 5**. The structural dissimilarity of the

linkers combined with the biological activity of the fusion proteins of the invention underscores the flexibility in linker selection. Accordingly, the fusion proteins of the present invention are not limited to cyclin-CDK fusion proteins containing the linkers which are specifically exemplified.

The fusion protein of **SEQ ID NO:7** has the additional features discussed above for allowing great flexibility in choice of purification schemes. The preferred aspect of this embodiment of the present invention is the segment of the molecule comprising the biologically active human cyclin D1-linker-human CDK4 sequence. This preferred sequence is set forth below as **SEQ ID NO:8.**

```
47                                                    MEHQ

51   LLCCEVETIR RAYPDANLLN DRVLRAMLKA EETCAPSVSY FKCVQKEVLP

101  SMRKIVATWM LEVCEEQKCE EEVFPLAMNY LDRFLSLEPV KKSRLQLLGA

151  TCMFVASKMK ETIPLTAEKL CIYTDNSIRP EELLQMELLL VNKLKWNLAA

201  MTPHDFIEHF LSKMPEAEEN KQIIRKHAQT FVALCATDVK FISNPPSMVA

251  AGSVVAAVQG LNLRSPNNFL SYYRLTRFLS RVIKCDPDCL RACQEQIEAL

301  LESSLRQAQQ NMDPKAAEEE EEEEEEEEVD LACTPTDVRD VDIASMGGGS

351  GGGSGGGSGG GSGGGSGGGS GLSSKGGGGS GGGGSGGGGS MATSRYEPVA

401  EIGVGAYGTV YKARDPHSGH FVALKSVRVP NGGGGGGGLP ISTVREVALL

451  RRLEAFEHPN VVRLMDVCAT SRTDREIKVT LVFEHVDQDL RTYLDKAPPP

501  GLPAETIKDL MRQFLRGLDF LHANCIVHRD LKPENILVTS GGTVKLADFG

551  LARIYSYQMA LTPVVVTLWY RAPEVLLQST YATPVDMWSV GCIFAEMFRR
```

601 KPLFCGNSEA DQLGKIFDLI GLPPEDDWPR DVSLPRGAFP PRGPRPVQSV

651 VPEMEESGAQ LLLEMLTFNP HKRISAFRAL QHSYLHKDEG NPE

Skilled artisans will recognize that the proteins of the present invention can be synthesized by a number of different methods. All of the amino acid compounds of the invention can be made by chemical methods well known in the art, including solid phase peptide synthesis, or recombinant methods. Both methods are described in U.S. Patent 4,617,149, herein incorporated by reference.

The principles of solid phase chemical synthesis of polypeptides are well known in the art and may be found in general texts in the area. See, e.g., H. Dugas and C. Penney, BIOORGANIC CHEMISTRY, (1981) Springer-Verlag, New York, pgs. 54-92. For examples, peptides may be synthesized by solid-phase methodology utilizing an Applied Biosystems 430A peptide synthesizer (commercially available from Applied Biosystems, Foster City California) and synthesis cycles supplied by Applied Biosystems. Protected amino acids, such as *t*-butoxycarbonyl-protected amino acids, and other reagents are commercially available from many chemical supply houses.

Sequential *t*-butoxycarbonyl chemistry using double couple protocols are applied to the starting *p*-methyl benzhydryl amine resins for the production of C-terminal carboxamides. For the production of C-terminal acids , the corresponding pyridine-2-aldoxime methiodide resin is used. Asparagine, glutamine, and arginine are coupled using preformed hydroxy benzotriazole esters. The following side chain protection may be used:

Arg, Tosyl
Asp, cyclohexyl
Glu, cyclohexyl
Ser, Benzyl
Thr, Benzyl
Tyr, 4-bromo carbobenzoxy

Removal of the *t*-butoxycarbonyl moiety (deprotection) may be accomplished with trifluoroacetic acid (TFA) in methylene chloride. Following completion of the synthesis the peptides may be deprotected and cleaved from the resin with anhydrous hydrogen fluoride containing 10% meta-cresol. Cleavage of the side chain protecting group(s) and of the peptide from the resin is carried out at zero degrees centigrade or below, preferably -20°C for thirty minutes followed by thirty minutes at 0°C.

After removal of the hydrogen fluoride, the peptide/resin is washed with ether, and the peptide extracted with glacial acetic acid and then lyophilized. Purification is accomplished by size-exclusion chromatography on a Sephadex G-10 (Pharmacia) column in 10% acetic acid.

The proteins of the present invention may also be produced by recombinant methods. Recombinant methods are preferred if a high yield is desired. A general method for the construction of any desired DNA sequence is provided in J. Brown, et al., Methods in Enzymology, 68:109 (1979). See also, J. Sambrook, et al., supra.

The basic steps in the recombinant production of desired proteins are:

a) construction of a synthetic or semi-synthetic DNA encoding the protein of interest;

b) integrating said DNA into an expression vector in a manner suitable for the expression of the protein of interest, either alone or as a fusion protein;

c) transforming an appropriate eukaryotic or prokaryotic host cell with said expression vector,

d) culturing said transformed or transfected host cell in a manner to express the protein of interest; and

e) recovering and purifying the recombinantly produced protein of interest.

In general, prokaryotes are used for cloning of DNA sequences in constructing the vectors of this invention. Prokaryotes may also be employed in the production of the protein of interest. For example, the Escherichia coli K12 strain 294 (ATCC No. 31446) is particularly useful for the prokaryotic expression of foreign proteins. A commercially available E. coli strain which is preferred for prokaryotic expression of the fusion proteins of the invention is designated DH10B. DH10B is available from Gibco BRL, P.O. Box 68, Grand Island, N.Y. 14072-0068. Other strains of E. coli which may be used (and their relevant genotypes) include the following.

Strain	**Genotype**
DH5a	F^- (ϕ80dlacZDM15), D(lacZYA-argF)U169 supE44, hsdR17(r_K^-, m_K^+), recA1, endA1, gyrA96, thi-1, relA1
HB101	supE44, hsdS20(r_B^- m_B^-), recA13, ara-14, proA2 lacY1, galK2, rpsL20, xyl-5, mtl-1, mcrB, mrr
JM109	recA1, e14$^-$(mcrA), supE44, endA1, hsdR17(r_K^-, m_K^+), gyrA96, relA1, thi-1, Δ(lac-proAB), F'[traD36, proAB+ $lacI^q$,lacZΔM15]
RR1	supE44, hsdS20(r_B^- m_B^-), ara-14 proA2, lacY1, galK2, rpsL20, xyl-5, mtl-5
chi1776	F^-, ton, A53, dapD8, minA1, supE42 (glnV42), D(gal-uvrB)40, minB2, rfb-2, gyrA25, thyA142, oms-2, metC65, oms-1, D(bioH-asd)29, cycB2, cycA1, hsdR2
294	endA, thi$^-$, hsr$^-$, hsm_K^+ (U.S. Patent 4,366,246)

LE392 F^-, hsdR514 (r^-m^-), supE44, supF58, lacY1, galK2, galT22, metB1, trpR55

These strains are all commercially available from suppliers such as: Bethesda Research Laboratories, Gaithersburg, Maryland 20877 and Stratagene Cloning Systems, La Jolla, California 92037; or are readily available to the public from sources such as the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland, 10852-1776.

Except where otherwise noted, these bacterial strains can be used interchangeably. The genotypes listed are illustrative of many of the desired characteristics for choosing a bacterial host and are not meant to limit the invention in any way. The genotype designations are in accordance with standard nomenclature. See, for example, J. Sambrook, et al., supra. A preferred strain of E. coli employed in the cloning and expression of the genes of this invention is RV308, which is available from the ATCC under accession number ATCC 31608, and is described in United States Patent 4,551,433, issued November 5, 1985. The three E. coli host cells transformed with the vectors described in Figures 1,2 and 3 and discussed in preceding sections will be publicly available upon issuance of a patent in a "Budapest Treaty" country and thus are the preferred means for prokaryotic expression of the fusion proteins which are described herein as illustrative of the fusion proteins of the invention. The fusion proteins produced by the E. coli "deposits" of the invention require solubilization, folding and phosphorylation for complete biological activity. While they are still preferred when substantial amounts of fusion protein are desired, the facile nature of numerous eukaryotic expression systems results in a preference for these systems when modest amounts of the biologically active fusion proteins are desired.

In addition to the strains of E. coli discussed supra, bacilli such as Bacillus subtilis, other enterobacteriaceae such as Salmonella typhimurium or Serratia marcescans, and various Pseudomonas species may also be used. In addition to these gram-negative bacteria, other bacteria, especially Streptomyces, spp., may be employed in the prokaryotic cloning and expression of the proteins of this invention.

Promoters suitable for use with prokaryotic hosts include the b-lactamase [vector pGX2907 (ATCC 39344) contains the replicon and b-lactamase gene] and lactose promoter systems [Chang et al., Nature (London), 275:615 (1978); and Goeddel et al., Nature (London), 281:544 (1979)], alkaline phosphatase, the tryptophan (trp) promoter system [vector pATH1 (ATCC 37695) is designed to facilitate expression of an open reading frame as a trpE fusion protein under control of the trp promoter] and hybrid promoters such as the tac promoter (isolatable from plasmid pDR540 ATCC-37282). However, other functional bacterial promoters, whose nucleotide sequences are generally known, enable one of skill in the art to ligate them to DNA encoding the proteins of the instant invention using linkers or adapters to supply any required restriction sites. Promoters for use in bacterial systems will also contain a Shine-Dalgarno sequence operably linked to the DNA encoding the desired polypeptides. These examples are illustrative rather than limiting.

The proteins of this invention may be synthesized either by direct expression or as a fusion protein comprising the protein of interest as a translational fusion with another protein or peptide which may be removable by enzymatic or chemical cleavage. It is often observed in the production of certain peptides in recombinant systems that expression as a fusion protein prolongs the lifespan, increases the yield of the desired peptide, or provides a convenient means of purifying the protein of interest. A variety of peptidases (e.g. trypsin) which cleave a polypeptide at specific sites or digest the peptides from the

amino or carboxy termini (e.g. diaminopeptidase) of the peptide chain are known. Furthermore, particular chemicals (e.g. cyanogen bromide) will cleave a polypeptide chain at specific sites. The skilled artisan will appreciate the modifications necessary to the amino acid sequence (and synthetic or semi-synthetic coding sequence if recombinant means are employed) to incorporate site-specific internal cleavage sites. See e.g., P. Carter, "Site Specific Proteolysis of Fusion Proteins", Chapter 13 in PROTEIN PURIFICATION: FROM MOLECULAR MECHANISMS TO LARGE SCALE PROCESSES, American Chemical Society, Washington, D.C. (1990).

In addition to cloning and expressing the genes of interest in the prokaryotic systems discussed above, the proteins of the present invention may also be produced in eukaryotic systems. The present invention is not limited to use in a particular eukaryotic host cell. A variety of eukaryotic host cells are available from depositories such as the American Type Culture Collection (ATCC) and are suitable for use with the vectors of the present invention. The choice of a particular host cell depends to some extent on the particular expression vector used to drive expression of the cyclin-CDK fusion protein-encoding nucleic acids of the present invention. Exemplary host cells suitable for use in the present invention are listed in Table I

Table I

Host Cell	Origin	Source
HepG-2	Human Liver Hepatoblastoma	ATCC HB 8065
CV-1	African Green Monkey Kidney	ATCC CCL 70
LLC-MK2	Rhesus Monkey Kidney	ATCC CCL 7
3T3	Mouse Embryo Fibroblasts	ATCC CCL 92
CHO-K1	Chinese Hamster Ovary	ATCC CCL 61
HeLa	Human Cervix Epitheloid	ATCC CCL 2
RPMI8226	Human Myeloma	ATCC CCL 155
H4IIEC3	Rat Hepatoma	ATCC CCL 1600

C127I	Mouse Fibroblast	ATCC CCL 1616
293	Human Embyronal Kidney	ATCC CRL 1573
HS-Sultan	Human Plasma Cell Plasmocytoma	ATCC CCL 1484
BHK-21	Baby Hamster Kidney	ATCC CCL 10

A preferred eukaryotic cell line of use in expressing the fusion proteins of this invention is the widely available cell line AV12-664 (hereinafter "AV12"). This cell line is available from the American Type Culture Collection under the accession number ATCC CRL 9595. The AV12 cell line was constructed by injecting a Syrian hamster in the scruff of the neck with human adenovirus 12 and isolating cells from the resulting tumor.

A wide variety of vectors, some of which are discussed below, exists for the transformation of such mammalian host cells, but the specific vectors described herein are in no way intended to limit the scope of the present invention. The sequences encoding the illustrative fusion proteins of the invention are easily removed from the deposited E. coli strains by reference to the Figures for selection of the appropriate restriction endonucleases and inserted in any of the vectors described herein through routine purification, ligation and transfection techniques.

The pSV2-type vectors comprise segments of the simian virus 40 (SV40) genome that constitute a defined eukaryotic transcription unit-promoter, intervening sequence, and polyadenylation site. In the absence of the SV40 T antigen, the plasmid pSV2-type vectors transform mammalian and other eukaryotic host cells by integrating into the host cell chromosomal DNA. A large number of plasmid pSV2-type vectors have been constructed, such as plasmid pSV2-gpt, pSV2-neo, pSV2-dhfr, pSV2-hyg, and pSV2-b-globin, in which the SV40 promoter drives transcription of an inserted gene. These vectors are suitable for use with the coding sequences of the present invention and are widely available from

sources such as the ATCC or the Northern Regional Research Laboratory (NRRL), 1815 N. University Street, Peoria, Illinois, 61604.

The plasmid pSV2-dhfr (ATCC 37146) comprises a murine dihydrofolate reductase (dhfr) gene under the control of the SV40 early promoter. Under the appropriate conditions, the dhfr gene is known to be amplified, or copied, in the host chromosome. This amplification can result in the amplification of closely-associated DNA sequences and can, therefore, be used to increase production of a protein of interest. See, e.g., R. T. Schimke, Cell, 35:705-713 (1984).

Plasmids constructed for expression of the proteins of the present invention in mammalian and other eukaryotic host cells can utilize a wide variety of promoters. The present invention is in no way limited to the use of the particular promoters exemplified herein. Promoters such as the SV40 late promoter, promoters from eukaryotic genes, such as, for example, the estrogen-inducible chicken ovalbumin gene, the interferon genes, the gluco-corticoid-inducible tyrosine aminotransferase gene, and the thymidine kinase gene, and the major early and late adenovirus genes can be readily isolated and modified to express the genes of the present invention. Eukaryotic promoters can also be used in tandem to drive expression of a coding sequence of this invention. Furthermore, a large number of retroviruses are known that infect a wide range of eukaryotic host cells. The long terminal repeats in the retroviral DNA frequently encode functional promoters and, therefore, may be used to drive expression of the nucleic acids of the present invention.

Plasmid pRSVcat (ATCC 37152) comprises portions of a long terminal repeat of the Rous Sarcoma virus, a virus known to infect chickens and other host cells. This long terminal repeat contains a promoter which is suitable for use in the vectors of this invention. H. Gorman, et al., Proceedings of the National Academy of Sciences (USA), 79:6777 (1982). The plasmid pMSVi (NRRL B-15929) comprises

the long terminal repeats of the Murine Sarcoma virus, a virus known to infect mouse and other host cells. The mouse metallothionein promoter has also been well characterized for use in eukaryotic host cells and is suitable for use in the expression of the nucleic acids of the present invention. The mouse metallothionein promoter is present in the plasmid pdBPV-MMTneo (ATCC 37224) which can serve as the starting material of other plasmids of the present invention.

An especially useful expression vector system employs one of a series of vectors containing the BK enhancer, an enhancer derived from the BK virus, a human papovavirus. The most preferred such vector systems are those which employ not only the BK enhancer but also the adenovirus-2-early region 1A (E1A) gene product. The E1A gene product (actually, the E1A gene produces two products, which are collectively referred to herein as "the E1A gene product") is an immediate-early gene product of adenovirus, a large DNA virus.

A preferred eukaryotic expression vector employed in the present invention is the phd series of vectors which comprise a BK enhancer in tandem with the adenovirus late promoter to drive expression of useful products in eukaryotic host cells. The construction and method of using the phd plasmid, as well as related plasmids, are described in U.S. Patents 5,242,688, issued September 7, 1993, and 4,992,373, issued February 12, 1991, all of which are herein incorporated by reference. Escherichia coli K12 GM48 cells harboring the plasmid phd are available as part of the permanent stock collection of the Northern Regional Research Laboratory under accession number NRRL B-18525. The plasmid may be isolated from this culture using standard techniques.

The plasmid phd contains a unique BclI site which may be utilized for the insertion of the gene encoding the protein of interest. The skilled artisan understands that linkers or adapters may be employed in cloning the gene of interest into this BclI site. The phd series of plasmids functions most efficiently when introduced into a host cell

which produces the E1A gene product, cell lines such as AV12-664, 293 cells, and others, described supra.

Transformation of the mammalian cells can be performed by any of the known processes including, but not limited to, the protoplast fusion method, the calcium phosphate co-precipitation method, electroporation and the like. See, e.g., J. Sambrook, et al., supra, at 3:16.30-3:16.66.

Other routes of production are well known to skilled artisans. In addition to the plasmids discussed above, it is well known in the art that some viruses are also appropriate vectors. For example, the adenovirus, the adeno-associated virus, the vaccinia virus, the herpes virus, the baculovirus, and the rous sarcoma virus are useful. Such a method is described in U.S. Patent 4,775,624, herein incorporated by reference. Several alternate methods of expression are described in J. Sambrook, et al., supra, at 16.3-17.44.

In addition to prokaryotes and mammalian host cells, eukaryotic microbes such as yeast cultures may also be used. The imperfect fungus Saccharomyces cerevisiae, or common baker's yeast, is the most commonly used eukaryotic microorganism, although a number of other strains are commonly available. For expression in Saccharomyces sp., the plasmid YRp7 (ATCC-40053), for example, is commonly used. See, e.g., L. Stinchcomb, et al., Nature (London), 282:39 (1979); J. Kingsman et al., Gene, 7:141 (1979); S. Tschemper et al., Gene, 10:157 (1980). This plasmid already contains the trp gene which provides a selectable marker for a mutant strain of yeast lacking the ability to grow in tryptophan.

Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [found on plasmid pAP12BD (ATCC 53231) and described in U.S. Patent No. 4,935,350, issued June 19, 1990, herein incorporated by reference] or other glycolytic enzymes such as enolase [found on plasmid pAC1 (ATCC 39532)], glyceraldehyde-3-phosphate dehydrogenase [derived from

plasmid pHcGAPC1 (ATCC 57090, 57091)], hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase, as well as the alcohol dehydrogenase and pyruvate decarboxylase genes of Zymomonas mobilis (United States Patent No. 5,000,000 issued March 19, 1991, herein incorporated by reference).

Other yeast promoters, which are inducible promoters, having the additional advantage of their transcription being controllable by varying growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein [contained on plasmid vector pCL28XhoLHBPV (ATCC 39475) and described in United States Patent No. 4,840,896, herein incorporated by reference], glyceraldehyde 3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose [e.g. GAL1 found on plasmid pRY121 (ATCC 37658)] utilization. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., European Patent Publication No. 73,657A. Yeast enhancers such as the UAS Gal from Saccharomyces cerevisiae (found in conjuction with the CYC1 promoter on plasmid YEpsec--hI1beta ATCC 67024), also are advantageously used with yeast promoters.

Skilled artisans also recognize that some alterations of SEQ ID NO:2, 3, 5, 6, 7 or 8 will fail to change the function of the amino acid compound. For instance, some hydrophobic amino acids may be exchanged for other hydrophobic amino acids. Those altered amino acid compounds which confer substantially the same function in substantially the same manner as the exemplified amino acid compound are also encompassed within the present invention. Typically such conservative substitutions attempt to preserve the: (a) secondary or tertiary structure of the polypeptide backbone; (b) the charge or hydrophobicity of the residue; or

(c) the bulk of the side chain. Some examples of such conservative substitutions of amino acids, resulting in the production of proteins which are functional equivalents of the proteins of SEQ ID NO:2, 3, 5, 6, 7 or 8 are shown in Table II, infra.

Table II

Original Residue	Exemplary Substitutions
Ala	Ser, Gly
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro, Ala
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

These substitutions may be introduced into the protein in a variety of ways, such as during the chemical synthesis or by chemical modification of an amino acid side chain after the protein has been prepared.

Alterations of the protein having a sequence which corresponds to the sequences of SEQ ID NO:2, 3, 5, 7 or 8 may also be induced by alterations of the nucleic acid

compounds which encodes these proteins. These mutations of the nucleic acid compound may be generated by either random mutagenesis techniques, such as those techniques employing chemical mutagens, or by site-specific mutagenesis employing oligonucleotides. Those nucleic acid compounds which confer substantially the same function in substantially the same manner as the exemplified nucleic acid compounds are also encompassed within the present invention.

Other embodiments of the present invention are nucleic acid compounds which comprise isolated nucleic acid sequences which encode SEQ ID NO: 2, 3, 5, 7, and 8. As skilled artisans will recognize, the amino acid compounds of the invention can be encoded by a multitude of different nucleic acid sequences because most of the amino acids are encoded by more than one nucleic acid triplet due to the degeneracy of the amino acid code. Because these alternative nucleic acid sequences would encode the same amino acid sequences, the present invention further comprises these alternate nucleic acid sequences.

The genes encoding the DNA molecules of the present invention may be produced using synthetic methodology. This synthesis of nucleic acids is well known in the art. See, e.g., E.L. Brown, R. Belagaje, M.J. Ryan, and H.G. Khorana, Methods in Enzymology, 68:109-151 (1979). The DNA segments corresponding to the fusion proteins are generated using conventional DNA synthesizing apparatus such as the Applied Biosystems Model 380A or 380B DNA synthesizers (commercially available from Applied Biosystems, Inc., 850 Lincoln Center Drive, Foster City, CA 94404) which employ phosphoramidite chemistry. In the alternative, the more traditional phosphotriester chemistry may be employed to synthesize the nucleic acids of this invention. See, e.g., M.J. Gait, ed., OLIGONUCLEOTIDE SYNTHESIS, A PRACTICAL APPROACH, (1984).

The DNA sequences of the present invention may be designed to possess restriction endonuclease cleavage sites at either end of the transcript to facilitate isolation from and integration into expression and amplification plasmids.

The choice of restriction sites are chosen so as to properly orient the coding sequence with control sequences to achieve proper in-frame reading and expression of the molecule. A variety of other such cleavage sites may be incorporated depending on the particular plasmid constructs employed and may be generated by techniques well known in the art.

In an alternative methodology, the human cyclin and human CDK coding regions of the desired DNA sequences can be generated using the polymerase chain reaction as described in U.S. Patent No. 4,889,818, which is herein incorporated by reference.

The preferred expression systems for use in the present invention are the various Baculovirus systems. The pFastBac1 expression system, which is commercially available from the Life Technologies group of Gibco BRL Products as Catalog No. 10360-016. Life Technologies, P.O. Box 68, Grand Island, NY 14072, Telephone: 800 828 6686, is the preferred expression system when modest amounts of biologically active fusion proteins are desired. The Bac-To-Bac Baculovirus Expression System has been used for expression of the sequences of the present invention and this system is also available from Life Technologies (Catalog No. 10359-016). The present inventors elected to deposit the DNA sequences encoding the illustrative cyclin-CDK fusion proteins as components of prokaryotic, lac operon-regulated expression systems due to the ability of the *E. coli* systems to produce large amounts of the fusion proteins and the ease with which skilled artisans can excise the desired coding sequences from the *E. coli* systems and insert them into these commerically available Baculovirus expression systems to thereby achieve the preferred mode of expressing modest amounts of the illustrative fusion proteins.

Baculovirus expression systems are well known in the art and numerous scientific articles and "methods" books are available on the subject. The present inventors have found the Life Technologies technical literature to provide excellent guidance for producing products of interest via

Baculovirus expression. The preferred techniques for Baculovirus expression of the sequences of the present invention are those provided in the product literature. Minor variations such as linker construction and the like are considered in light of the advanced state of this art as too trivial to warrant discussion. In the event skilled artisans elect to depart from the commercially available Baculovirus systems, the present inventors recommend Baculovirus Expression Vectors-A Laboratory Manual, O'Reilly, David R., Miller, Lois K., and Luckow, Verne A., W. H. Freeman and Company, New York, New York as a source of additional information on any protocol required for successful expression of polypeptides in Baculovirus systems.

The assays which are greatly advantaged by the fusion proteins of the present invention are well illustrated in two recent scientific publications: Connell-Crowley, L., *et al.*, Mol. Biol. of the Cell **4**, 79-92 (1993) and Desai, D., Mol. Biol. of the Cell **3**, 571-582 (1992).

The examples provide sources for reagents, however it will be understood that numerous vendors market reagents of high quality for use in the protocols and procedures described below and the substitution of reagents or protocols is contemplated by the present invention and embraced in the scope thereof. All temperatures unless otherwise noted are expressed in degrees Centigrade. All percentages are on a weight per weight basis unless otherwise noted.

Skilled artisans wishing to practice the recombinant DNA aspects of the present invention are directed to the NIH guidelines for information on research involving recombinant DNA molecules. A copy of the current guidelines can be obtained from Office of Recombinant DNA Activities, National Institutes of Health, Building 31, Room 4B11, Bethesda, MD 20892. Compliance with all such current regulations regarding vector selection, expression of human and animal genes and containment requirements is required by law.

The examples are intended to further illustrate the present invention and are not to be interpreted as limiting on the scope thereof. While the examples and detailed description sections of the present invention are sufficient to guide anyone of ordinary skill in the art in the practice of the present invention, skilled artisans are also directed to *Molecular Cloning A Laboratory Manual* Second Edition, Sambrook,J., Fritsch, E. F., and Maniatis, T., Cold Spring Harbor Press 1989 and *Current Protocols In Molecular Biology*, Ausubel, F.M., Brent,R., Kingston,R.E., Moore, D. D., Seidman,J.G., Smith, J.A., and Struhl, K., Ed. Greene Publishing Associates and Wiley-Interscience 1989. The aforementioned resources provide an excellent technical supplement to any discourse in genetic engineering.

Example 1

Production of Baculovirus System for Expression of SEQ ID NO:2

A sample of NRRL B-21490 is obtained from the NRRL. The sample is cultured according to well known procedures using standard media containing Ampicillin for selection of the desired transformed phenotype.

Plasmid isolation is accomplished in accordance with standard methodology. See e.g. Sambrook and Maniatis, supra.

The desired fragment is excised from plasmid pK415 (See Figure 1) by sequential digestion with the restriction endonucleases, AscI and Sse 8387I. The AscI digestion is performed using New England Biolabs reagents and protocols. The restriction endonuclease Sse 8387I is available from Takara Biomedicals via PanVera Corp., 565 Science Drive, Madison, WI 53711 (1 800 791-1400). The vendors instructions on digestion procedures are recommended.

pFastBacI is digested with BssHII (New England Biolabs) and PstI (New England Biolabs) in accordance with vendors instructions and the large fragment is isolated. A restriction site and function map of pFastBac1 is provided at

page 5 of the GibcoBRL/Life Technologies Catalog Number 10359-016 (Instruction Manual-BAC-TO-BAC™ Baculovisur Expression System). The catalog is herein incorporated by reference. The fusion protein encoding sequence is then ligated into the pFastBac1 vector using standard ligation reagents and conditions. Preferred ligation reagents and conditions are set forth at pages 7 and 8, Section 3.3, of GibcoBRL/Life Technogies Catalog Number 10359-016. Page 5 of GibcoBRL/Life Technogies Catalog Number 10359-016 provides DNA sequence information and restriction endonuclease cleavage sites for the multiple cloning site of pFastBac1 and is therefore useful in the event skilled artisans elect to fragment the sequence from p415 or excise it by other than the restriction endonucleases suggested above and utilize linkers to facilitate the subsequent ligation into pFastBac1.

Transposition of the pFastBac1 vector comprising the fusion protein of plasmid pK415 into DH10Bac10 (competent cells are provided as part of the expression kit accompanying pFastBac1 in Catalog Number 10359-16) is conducted in accordance with the teachings of page 8 of GibcoBRL/Life Technogies Catalog Number 10359-016.

Isolation of Recombinant Bacmid DNA is accomplished in accordance with the teachings of pages 8 and 9 of GibcoBRL/Life Technogies Catalog Number 10359-016.

Transfection of Sf9 cells with recombinant Bacmid DNA, harvesting and storage of the recombinant Baculovirus, and Infection of Insect Cells with recombinant Baculovirus particles is accomplished with the teachings at pages 9 and 10 of GibcoBRL/Life Technogies Catalog Number 10359-016.

Example 2

Production of Baculovirus System for Expression of SEQ ID NO:4

Baculovirus expression systems were constructed in substantial accordance with the teachings of Example 1. Plasmid pK480 from E. coli/pK485 was used in place of plasmid

pk415 as the source of the DNA sequence encoding the fusion protein of interest.

Example 3

Production of Baculovirus System for Expression of SEQ ID NO:6

Baculovirus expression systems were constructed in substantial accordance with the teachings of Example 1. Plasmid pK485 from E. coli/pK480, NRRL number B21491, was used in place of plasmid pK415 as the source of the DNA sequence encoding the fusion protein of interest. With the exception of the substitution of plasmid pK480 for plasmid pK415 all steps of this Example 3 were carried out in conformance with the teachings of Example 1.

Example 4

Purification of Co-expressed D1.K4

Affinity chromatography resins for fusion protein purification are readily constructed from commercially available reagents using techniques well known in the art.

CNBr-activated Sepharose 4B (Pharmacia Fine Chemicals) is the preferred matrix for linkage of appropriate monoclonal or polyclonal antibodies to allow antibody-based affinity purification of the fusion proteins. Pharmacia Fine Chemicals publishes "Affinity Chromatography-Principles and Methods". This manual sets forth all steps in preparing the affinity resin and performing the antibody-based affinity purification steps. The manual is available from Pharmacia Fine Chemicals, Box 175, S-751 04 Uppsala 1, Sweden.

Example 5

Strepavadin Purification of Cyclin-CDK Fusion Proteins

The SF9 cells which were utilized in Examples 1-3 as the host cells for Baculovirus expression were collected by centrifugation and resuspended and lysed via sonication at 4°C in Resuspension Buffer at a density of 8 X 10^6/mL.

Resuspension buffer is 50mM HEPES pH 7.5, 0.32M Sucrose, 0.1 mM PMSF, 1.0mM DTT, 1mM EDTA and 80mM ß-glycerophosphate.

500µL of the SF9 extract was added to 200µL of Streptavidin Paramagnetic Beads (Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399) and the mixture was incubated at room temperature for 45 minutes. The paramagnetic beads were pelleted at room temperature using a MagneSphere Technology Magnetic separation stand (Promega). The beads were washed three times with 1 mL of 1XPBS/25 mg/ml BSA (or 0.1% Tween 20) at room temperature.

The fusion protein was eluted from the beads in 120 µL of Elution Buffer A for 30 minutes at room temperature. Elution Buffer A is 25mM HEPES pH 7.5, 0.1 mM PMSF, 1mM d-Biotin 0.1mM DTT, 20mM ß-glycerophosphate, 1mMNaF, 10mM Sodium Orthovanadate and 10% glycerol.

The purified fusion protein was stored at -70° C until ready for use.

Example 6

Ni-NTA Purification of Cyclin-CDK Fusion Proteins

8 X 10^6 SF6 cells/mL (from Examples 1-3) were collected by centrifugation and resuspended and lysed at 4°C in Resuspension Buffer. 1.0 mL of the insect cell extract was added to 3.0 mL of Ni-NTA agarose (Qiagen Inc., 9259 Eton Avenue, Chatsworth, CA 91311), which was previously equilibrated with Wash Buffer. Wash Buffer is 50mM HEPES pH 7.5, 300 mM NaCl, 20mM Imidizole and 0.1 mM PMSF.

The extract agarose mixture was incubated at 4°C for 4 hours. The mixture was gently agitated during the incubation. The agarose was then pelleted by centrifugation at 2000xg for two minutes and then washed three times with 5.0mL of 1XPBS at 4°C with agitation. The fusion protein was eluted from the agarose in 750 µL of Elution Buffer B for 1 hour at 4°C with agitation. Elution Buffer B is 50mM HEPES pH 7.5, 300mM NaCl, 250mM Imidizole, 0.1 mM PMSF, 10mM Sodium Orthovanadate, 1mM NaF and 20mM ß-glycerophosphate. The

eluted fusion protein was dialyzed in 3.0L of Dialysis Buffer overnight at 4°C. Dialysis Buffer is 25mM HEPES ph 7.5, 10% glycerol, 0.01% Triton-X, 0.1mM PMSF,20mM ß-glycerophosphate, 1mM NaF and 10mM Sodium Orthovanadate.

The dialyzed fusion protein was stored at -70°C.

Example 7

Purification of Co-expressed D1.K4 Individual Units

Purification of co-expressed cyclin D1 and cdk4 was performed at Spinx Pharmaceuticals. Insect cell pellets were homogenized at 1:10 in 50 mM HEPES pH 7.5, 320 mM Sucrose, 1mM DTT, 0.1mM PMSF, 1mM EGTA, 1mM EDTA and 20µg/ml leupeptin. The lysed cells were spun for 1.5 hrs. at 100,000 xg to remove cytosol then equilibrated a Poros Q column in Equilibration Buffer (25mM Tris pH 8.0, 10% glycerol, 1mM DTT, 0.1mM PMSF, 1mM EDTA, and 20 µg/ml leupeptin). The lysates were loaded onto a Poros Q column at 5ml/L of infected insect cells. The Poros Q column was washed with 10-column volumes of Equilibration buffer. The column was eluted with 0-1M NaCl gradient collecting 2ml/fraction. The column fractions were assayed for activity and peak fractions were pooled. The resulting pool was diluted to give a final NaCl concentration of 100mM. The dilute pool fractions were loaded onto a Hydroxapatite column equilibrated with 25mM Tris pH 8.0, 0.1 mM PMSF, 1mM EDTA, and 20 µg/ml leupeptin. The Hydroxapatite column was washed with 10-column volumes of Equilibration buffer and eluted cyclin D1 and cdk4 with 0-400mM potassium phosphate, pH 7.5. Column fractions were assayed for activity and the peak fractions pooled. The eluted protein was stored at -70C.

Example 8

Immunoprecipitation of D1.K4 Fusion

5×10^6 cells/mL were lysed in IP Lysis Buffer on ice for 30 minutes (IP Lysis Buffer: 50mM HEPES pH 7.4, 150 mM NaCl, 1 mM EDTA, 1mM DTT, 2.5mM EGTA, 0.1% Tween 20, 10%

Glycerol, 0.1mM PMSF, 500µM ATP, 10mM ß-glycerophosphate, 1mM NaF, and 0.1mM orthovanadate). The cells were sonicated three times on ice for 10 seconds each time, and the lysates were clarified for 5 minutes at 10,000 rpm and 4°C. 20µL of myc antibody (100µg/mL commercially available from Oncogene Science, Cambridge, Mass.) was added to 500µL of clarified cell lysate. The mixture was incubated with agitation for 3 hours at 4°C. 50µl of 50% Protein-G Agarose (Boehringer Mannheim), which had been washed with IP Lysis Buffer, was then added to each sample. The samples were incubated with agitation for 2-5 hours at 4°C. The Protein-G-Agarose was pelleted and washed 4X with IP Lysis Buffer and then 2X with 50mM HEPES pH 7.4 and 1mM DTT. The washed Protein-G-Agarose was resuspended in Kinase Reaction Buffer.

Example 9

Assays for Clyclin D1 and cdk4

Partially purified co-expressed or fused cyclin D1 and cdk4 were assayed for Rb kinase activity. Co-expressed cyclin D1 and cdk4 were partially purified as described above. Fused cyclin D1-cdk4 was partially purified by streptavidin beads, Ni-NTA agarose, and by immunoprecipitation. In immunoprecipitations, fused cyclin D1-cdk4 expressed in stably transfected Rat Embryo Fibroblasts (E3600NA-FPr-5) were partially purified as described in Matsushime et al., 1994. Kinase reactions with various amounts of partially purified cyclin D1 and cdk4 from insect cells contained: 50mM HEPES pH 7.5, 10mM $MgCl_2$, 0.2µCi [gamma-^{32} P]ATP (Amersham, 6,000 Ci/mmol), 0.12µg pRb (full-length protein from Immuno Pharmaceutics), 0.1mM sodium orthovanadate, 10mM ß-glycerophophate and 1mM NaF in a total of 100µL. Kinase reactions with immunoprecipitated fusion protein on Protein-G-Agarose (Boehringer Mannheim) from the REF cell line were resuspended in 50 µl of Kinase Reaction Buffer (50mM HEPES pH 7.5, 10mM $MgCl_2$, 10.0µCi [gamma-^{32}p]ATP (Amersham, 6,000 Ci/mmol), 0.2µg pRb (full-length protein from Immuno Pharmaceutics), 1mM DTT, 2.5 mM EGTA, 20µM ATP,

0.1mM sodium orthovanadate, 10mM ß-clycerophophate and 1mM NaF). Reactions were incubated at 30°C for 30 minutes, boiled for 5 minutes, and half of the reaction was loaded onto a 12.5% SDS-polyarcylamide gel. The gel was transferred to Hybond-ECL nitrocellulose (Amersham) and exposed to Hyperfilm-ECL (Amersham).

Example 10

Immunoblots

For protein detection of cyclin D1 and cdk4, nitrocellulose membranes were blocked with 5% dry milk in 1 x PBS for 30 to 60 minutes. Membranes were washed 3x, 10 minutes for each wash, in 1x PBS/0.1% Tween 20. The membrane was incubated with primary antibody (cyclin D1 or cdk4) at a 1:2000 dilution in 1X PBS/0.1% Tween 20/1% Milk for 1 hour at room temperature then washed 3X for 10 minutes each in 1X PBS/0.1% Tween 20. The membrane was then incubated with a secondary antibody (horse radish peroxidase conjugated goat anti-mouse or rabbit antibody from Amersham) at a 1:1000 dilution in 1X PBS/0.1% Tween 20/1% Milk for 25 minutes at room temperature. The membrane was washed 6X in PBS/0.1% Tween 20, 2X in 1X PBS, and developed with Amersham ECL detection reagents. The results indicated that the fusion protein had substantially the same amount of activity as the individual subunits.

We Claim:

1. A fusion protein comprising a human cyclin and a human CDK.

2. The fusion protein of Claim 1 wherein one or more of the amino acid residues of said human cyclin are replaced by conservative substituions.

3. The fusion protein of Claim 1 wherein one or more of the amino acid residues of said human CDK are replaced by conservative amino acid substituions.

4. The fusion protein of Claim 1 wherein said human cyclin in human cyclin D1.

5. The fusion protein of Claim 1 wherein said human CDK is human CDK4.

6. The fusion protein of Claim 1 that is SEQ ID NO:2.

7. The fusion protein of Claim 1 that is SEQ ID NO:3.9.

8. The fusion protein of Claim 1 that is SEQ ID NO:5.

9. The fusion protein of Claim 1 that is SEQ ID No. 7.

FIG. 1

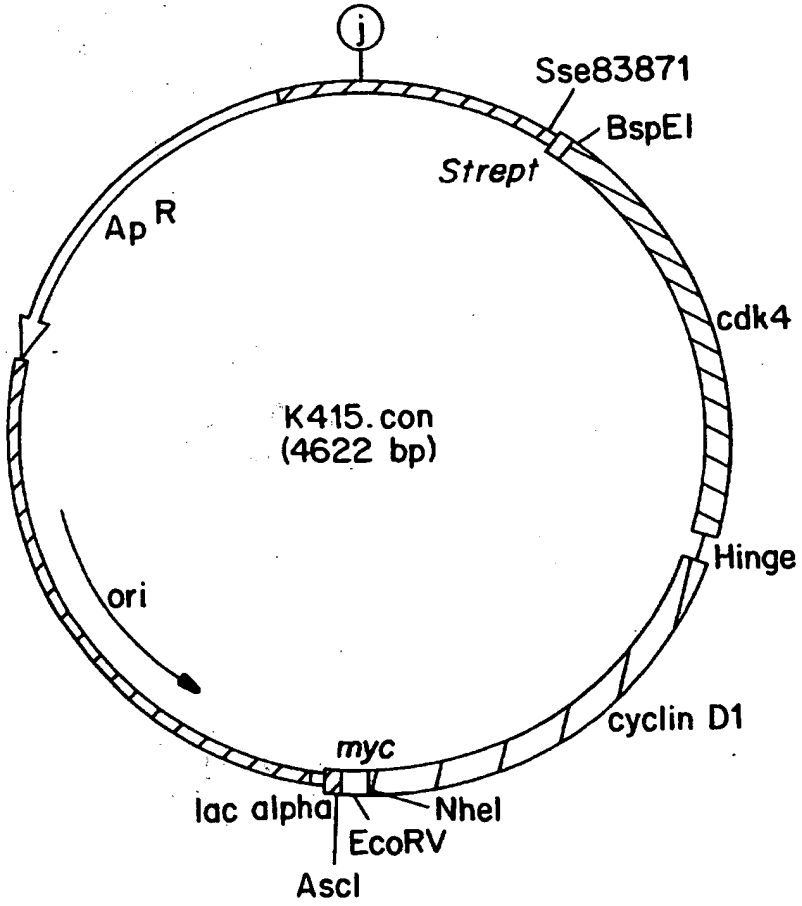

Sse83871
BspEI
Strept
Ap R
cdk4
K415.con
(4622 bp)
Hinge
ori
cyclin D1
myc
lac alpha
NheI
EcoRV
AscI

FIG. 2

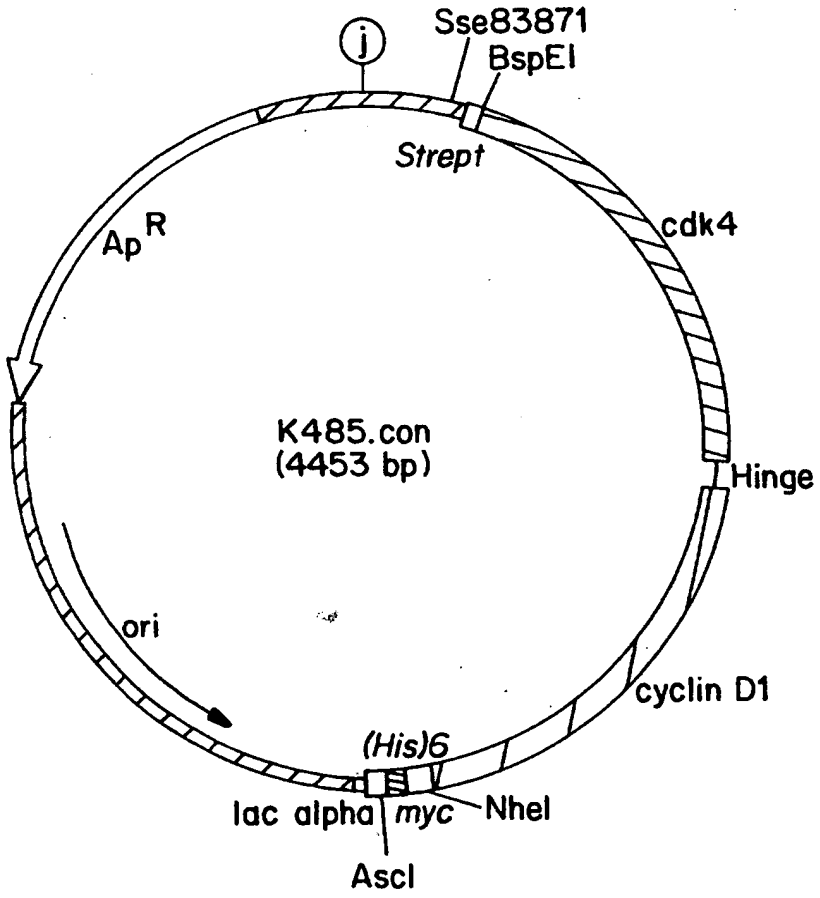
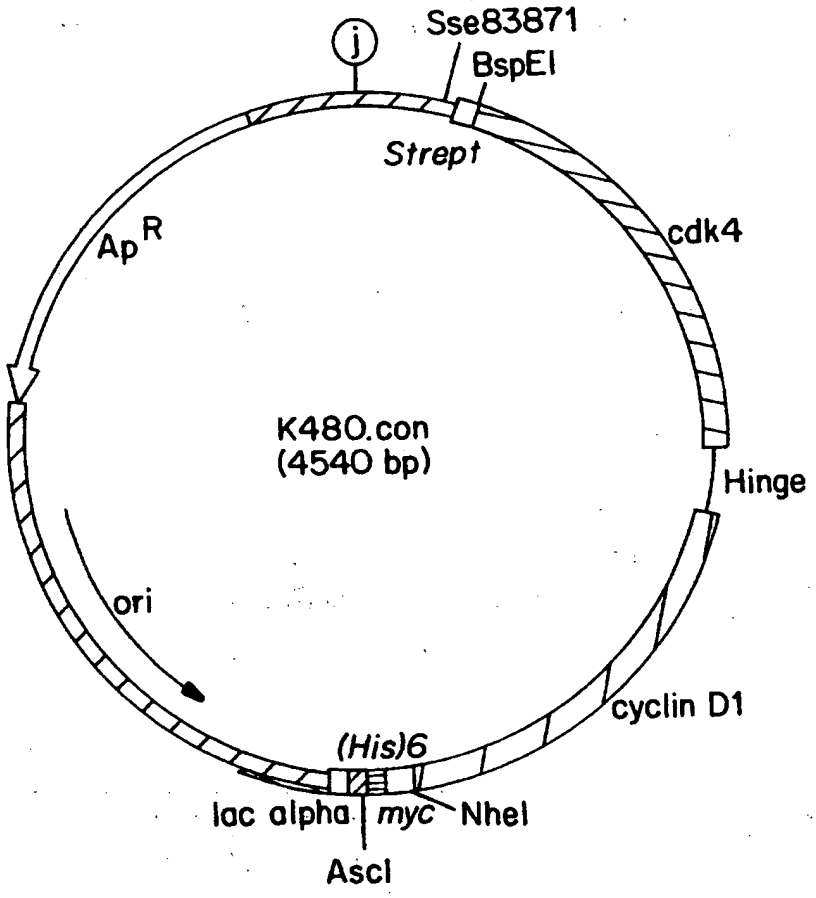

Sse83871
BspEI
Strept
cdk4
Ap R
K485.con
(4453 bp)
Hinge
ori
cyclin D1
(His)6
lac alpha
myc
NheI
AscI

FIG. 3


Sse83871
BspEI
Strept
cdk4
Ap R
K480.con
(4540 bp)
Hinge
ori
cyclin D1
(His)6
lac alpha
myc
NheI
AscI

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 9 and 41, line 11 and 19.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet [XX]

Name of depositary institution: NRRL, Northern Regional Research Laboratory

Address of depositary institution *(including postal code and country)*:

1815 North University Street
Peoria, Illinois 61604
US

Date of deposit	Accession Number
09 August 1995 (09.08.95)	B-21490

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet [ ]

Escherichia coli
(K415: cyclin D1/cdk4 fusion)

This microorganism identified above was accompanied by:
a proposed taxonomic designation

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
[ ] This sheet was received with the international application	[ ] This sheet was received by the International Bureau on: 27 MARCH 1997
Authorized officer	Authorized officer (M. Fourné-Godbersen)

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description
on page 17 . line 3

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet XX

Name of depositary institution

NRRL, Northern Regional Research Laboratory

Address of depositary institution *(including postal code and country)*

1815 North University Street
Peoria, Illinois 61604
US

Date of deposit	Accession Number
09 August 1995 (09.08.95)	B-21492

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

Escherichia coli
(K485 cyclin D1/cdk4 fusion)

The microorganism identified above was accompanied by:
a proposed taxonomic designation

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☐ This sheet was received with the international application	☐ This sheet was received by the International Bureau on: 27 MARCH 1997
Authorized officer	Authorized officer (M. Fourné-Godbersen)

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 24 and 43, line 5 and 9

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet [X]

Name of depositary institution

NRRL, Northern Regional Research Laboratory

Address of depositary institution *(including postal code and country)*

1815 North University Street
Peoria, Illinois 61604
US

Date of deposit	Accession Number
09 August 1995 (09.08.95)	B-21491

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet [ ]

Escherichia coli
(K480: cyclin D1/cdk4 fusion)

The microorganism identified above was accompanied by:
a proposed taxonomic designation

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
[ ] This sheet was received with the international application	[ ] This sheet was received by the International Bureau on: 27 MARCH 1997
Authorized officer	Authorized officer (M. Fourné-Godbersen)

Form PCT/RO/134 (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/00140

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07K 14/00; C12P 21/02; C12N 9/12; C07H 21/04
US CL : 530/350; 435/69.7, 194; 536/23.4

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 435/69.7, 194; 536/23.4

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN-indices Bioscience and Patents; files CJACS and USPATFULL
search terms: cyclin, D1, cdk4, fusion, fused, chimeric, chimaeric

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,338,669 A (GILLIES et al.) 16 August 1994, col. 1, lines 10-35.	1-5
Y	EWEN, M.E. et al. Functional Interactions of the Retinoblastoma Protein with Mammalian D-type Cyclins. Cell. 07 May 1993, Vol.73, pages 487-497, especially page 489.	1-5

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 11 APRIL 1997	Date of mailing of the international search report 05 MAY 1997
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer KAWAI LAU Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*